



3FW

Attorney Docket No.: 2543-1-034

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS : Jonathan A. Terrett
SERIAL NO. : 10/736,461
FILED : December 15, 2003
FOR : CANCER ASSOCIATED PROTEIN

Certificate of Mailing Under 37 CFR 1.8

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to COMMISSIONER FOR PATENTS, P.O. Box 1450, Alexandria, VA 22313-1450 on May 17, 2007.

Carolyn Di Meglio
(Name of Depositor)

Carolyn Di Meglio 5/17/07
(Signature and Date)

PETITION FOR GRANT OF PRIORITY UNDER 35 USC 119

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

Applicant hereby petitions for grant of priority of the present Application on the basis of the following prior filed foreign Application:

<u>COUNTRY</u>	<u>SERIAL NO.</u>	<u>FILING DATE</u>
Great Britain	GB0114644.8	JUNE 15, 2001

To perfect Applicant's claim to priority, a certified copy of the above listed prior filed Application is enclosed.

Acknowledgment of Applicant's perfection of claim to priority is accordingly requested.

Respectfully submitted,

Sarah J. Fashena
Sarah J. Fashena, Ph.D.
Agent for Applicant
Registration No. 57,600

KLAUBER & JACKSON LLC
411 Hackensack Avenue
Hackensack, NJ 07601
(201)487-5800
Date: May 17, 2007

Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (2) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with patent application GB0114644.8 filed on 1 June 2001.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., llo, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.

Signed



Dated 1 May 2007

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP9 1RH

15 JUN 2001

1. Your reference: **NJL/P32638GB**

18JUN01 E637699-3 D00056
E01/7700 0 00-0114644.8

2. Patent application number
(The Patent Office will fill in this part)

0114644.8

3. Full name, address and postcode of the or of each applicant (underline all surnames)

Oxford GlycoSciences (UK) Limited
The Forum
86 Milton Park
Abingdon
Oxfordshire
OX14 4RY
United Kingdom

Patents ADP number (if you know it)

If the applicant is a corporate body, give the country/state of its incorporation

7112386002

4. Title of the invention

Protein

5. Name of your agent (if you have one)

"Address for service" in the United Kingdom to which all correspondence should be sent (including the postcode)

Kilburn & Strode
20 Red Lion Street
London
WC1R 4PJ

Patents ADP number (if you know it)

125001

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (if you know it) the or each application number

Country

Priority application number
(if you know it)

Date of filing
(day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (Answer 'Yes' if:

- a) any applicant named in part 3 is not an inventor, or
- b) there is an inventor who is not named as an applicant, or
- c) any named applicant is a corporate body.

See note (d))

YES

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form

Description	49
Claim(s)	4
Abstract	1
Drawing(s)	6 + 6

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (*Patents Form 7/77*)

Request for preliminary examination and search (*Patents Form 9/77*)

Request for substantive examination (*Patents Form 10/77*)

Any other documents
(please specify)

11. I/We request the grant of a patent on the basis of this application.

Signature

Date 15 June 2001

12. Name and daytime telephone number of person to contact in the United Kingdom

Nicholas J Lee

Tel: 020 7539 4200

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 0645 500505.
- Write your answers in capital letters using black ink or you may type them.
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.
- Once you have filled in the form you must remember to sign and date it.
- For details of the fee and ways to pay please contact the Patent Office.

PROTEIN

The present invention relates to a sulfate transporter protein and its involvement in human cancer, particularly breast and colorectal cancer, compositions comprising the protein, including vaccines and antibodies which are immunospecific for the protein.

Breast cancer is the most frequently diagnosed cancer in women. The implementation of screening programs for the early detection of breast cancer, and the advent of anticancer treatments, such as chemotherapy, radiotherapy and anti-oestrogen therapies, to augment surgical resection have improved the survival of breast cancer patients.

However, some breast tumours become refractory to such treatments, as the cancer cells develop resistance to chemotherapy drugs or lose their hormone sensitivity, leading to recurrent or metastatic disease which is often incurable. More recently, attention has focussed on the development of immunological therapies (Green, et al. *Cancer Treat. Rev.* **26**, 269-286 (2000); Davis, I.D. *Immunol. Cell Biol.* **78**, 179-195 (2000); Knuth, et al. *Cancer Chemother Pharmacol.* **46**, S46-51 (2000); Shiku, et al. *Cancer Chemother. Pharmacol.* **46**, S77-82 (2000); Saffran, et al. *Cancer Metastasis Rev.* **18**, 437-449 (1999)), such as cancer vaccines and monoclonal antibodies (mAbs), as a means of initiating and targeting a host immune response against tumour cells. In 1998 the FDA approved the use of Herceptin (Stebbing, et al. *Cancer Treat. Rev.* **26**, 287-290 (2000); Dillman, et al. *Cancer Biother. Radiopharm.* **14**, 5-10 (1999); Miller, et al. *Invest. New Drugs* **17**, 417-427 (1999)), a mAb that recognises the erbB2/HER2-neu receptor protein, as a treatment for metastatic breast cancer. In combination with chemotherapy, Herceptin has been shown to prolong the time to disease progression, when compared to patients receiving chemotherapy alone (Baselga, et al. *Cancer Res.* **58**, 2825-2831 (1998)). Herceptin, however, is only effective in treating the 10-20% of patients whose tumours

over-express the erbB2 protein. Thus, the identification of other suitable targets or antigens for immunotherapy of breast cancer has become increasingly important.

5 Excluding skin cancers, colorectal cancer is the third most common cancer diagnosed in men and women in the United States. The American Cancer Society (<http://www.cancer.org/>) estimates that about 98,200 new cases of colon cancer (46,200 men and 52,000 women) and 37,200 new cases of rectal cancer (21,100 men and 16,100 women) will be diagnosed in 2001. Colon cancer is expected to be responsible for about 48,100 deaths (23,000 men and 25,100 women) during 2001. About 8,600 people (4,700 men and 3,900 women) will die from rectal cancer during 2001. The 5-year relative survival rate is 90% for people whose colorectal cancer is found and treated in an early stage, before it has spread. But, only 37% of colorectal cancers are found at that early stage. Once the cancer has spread to nearby organs or lymph nodes, the 5-year relative survival rate goes down to 65%. For people whose colorectal cancer has spread to distant parts of the body such as the liver or lungs, the 5-year relative survival rate is 8%.

Carcinoembryonic antigen (CEA) and CA 19-9 are substances produced by cells of most colon and rectal cancers and released into the bloodstream. These markers, however, can be high for reasons other than cancer, or can be normal in a person who has cancer. Therefore a need exists for more specific markers of colorectal cancer.

Surgery is the main treatment for colorectal cancer. Radiation therapy is often used after surgery to kill unremoved deposits and to prevent local recurrences. Adjuvant chemotherapy may also be used, with drugs such as Fluorouracil (5-FU), optionally with leucovorin or levamisole, and Irinotecan (CPT-11).

There are no generally approved immunotherapeutic drugs for the treatment of colorectal cancer, despite the fact that immunotherapy may offer the greatest potential after surgical

resection in the adjuvant setting. Edrecolomab (monoclonal antibody 17-1A, or Panorex), however, is an adjunctive therapy for colorectal cancer which is in clinical trials in the UK and the US, and which has already been approved in Germany. Identification of new suitable targets or antigens for immunotherapy of colorectal cancer is therefore highly important.

An ideal protein target for cancer immunotherapy should have a restricted expression profile in normal tissues and be over-expressed in tumours, such that the immune response will be targeted to tumour cells and not against other organs. In addition, the protein target should be exposed on the cell surface, where it will be accessible to therapeutic agents. Tumour antigens have been identified for a number of cancer types, by using techniques such as differential screening of cDNA (Hubert, et al. *Proc. Natl. Acad. Sci. USA* **96**, 14523-14528 (1999); Lucas, et al. *Int. J. Cancer* **87**, 55-60 (2000)), and the purification of cell-surface antigens that are recognised by tumour-specific antibodies (Catimel, et al. *J. Biol. Chem.* **271**, 25664-25670 (1996)). We have identified a protein, designated DTD, which shows restricted expression to a few tissues, with elevated expression in breast and colorectal tumours, suggesting that it may be a suitable target for cancer therapy and diagnosis.

The Diastrophic Dysplasia (DTD) gene was cloned in 1994 (Hastbacka J, et al. *Cell* 1994 Sep 23;78(6):1073-87). Its sequence is stored in the Swiss-Prot database (held by the Swiss Institute of Bioinformatics (SIB) and available at <http://www.expasy.ch/>) under the accession number P50443 and encodes a putative 12-membrane sulfate transporter of 739 amino acids long (Figure 1). It is located on chromosome 5q31-q34.

Mutations in DTD constitute a family of recessively inherited osteochondrodysplasias including Diastrophic Dysplasia (Hastbacka et al., *supra*), Achondrogenesis Type IB (Superti-Furga A, et al. *Nat Genet* 1996 Jan;12(1):100-2) and Atelosteogenesis Type II

(Hastbacka J, et al. *Am J Hum Genet* 1996 Feb;58(2):255-62). JP11146790 disclosed a novel vector for the expression of DTD and the use thereof to screen for agents treating a human bone/cartilage disease, as well as the use of vector in the preparation of a drug for the treatment of human bone/cartilage diseases.

5

Thus, in a first aspect, the present invention provides a method of screening for and/or diagnosis of cancer in a subject, which method comprises the step of detecting and/or quantifying the amount of a polypeptide in a biological sample obtained from said subject, wherein the polypeptide:

10

- a) comprises or consists of the amino acid sequence shown in figure 1;
- b) is a derivative having one or more amino acid substitutions, deletions or insertions relative to the amino acid sequence shown in figure 1; or
- c) is a fragment of a polypeptide as defined in a) or b) above, which is at least ten amino acids long.

15

In a second aspect, the present invention provides a method for the prophylaxis and/or treatment of cancer in a subject, which comprises administering to said subject a therapeutically effective amount of at least one polypeptide as defined in the first aspect of the invention.

20

In a third aspect, the present invention provides the use of at least one polypeptide as defined in the first aspect of the invention in the preparation of a medicament for use in the prophylaxis and/or treatment of cancer.

25

The present invention finds use in the treatment, prophylaxis and diagnosis of breast cancer and colorectal cancer. As used herein, breast tissue refers to the breast itself as well as the tissue adjacent and/or within the strata underlying the breast. The subject may be a

mammal and is preferably a human, although monkeys, apes, cats, dogs, cows, horses and rabbits are within the scope of the present invention.

The polypeptides of the invention or fragments thereof may be provided in isolated or recombinant form, and may be fused to other moieties. In particular, fusions of the polypeptides or fragments thereof with localisation-reporter proteins such as the Green Fluorescent Protein (U.S. Patent Nos. 5,625,048, 5,777,079, 6,054,321 and 5,804,387) or the DsRed fluorescent protein (Matz,*et al* (1999) *Nature Biotech.* 17:969–973) are specifically contemplated. The polypeptides or fragments thereof may be provided in substantially pure form, that is to say free, to a substantial extent, from other proteins. Thus, a polypeptide may be provided in a composition in which it is the predominant component present (i.e. it is present at a level of at least 50%; preferably at least 75%, at least 90%, or at least 95%; when determined on a weight/weight basis excluding solvents or carriers).

In order to more fully appreciate the present invention, polypeptides within the scope of a)-c) above will now be discussed in greater detail.

Polypeptides within the scope of a)

A polypeptide within the scope of a), may consist of the particular amino acid sequence given in Figure 1 or may have an additional N-terminal and/or an additional C-terminal amino acid sequence relative to the sequence given in Figure 1.

Additional N-terminal or C-terminal sequences may be provided for various reasons. Techniques for providing such additional sequences are well known in the art. Additional sequences may be provided in order to alter the characteristics of a particular polypeptide. This can be useful in improving expression or regulation of expression in particular expression systems. For example, an additional sequence may provide some

protection against proteolytic cleavage. This has been done for the hormone Somatostatin by fusing it at its N-terminus to part of the β galactosidase enzyme (Itakwa *et al.*, *Science* **198**: 105-63 (1977)).

- 5 Additional sequences can also be useful in altering the properties of a polypeptide to aid in identification or purification. For example, a fusion protein may be provided in which a polypeptide is linked to a moiety capable of being isolated by affinity chromatography. The moiety may be an antigen or an epitope and the affinity column may comprise immobilised antibodies or immobilised antibody fragments which bind to said antigen or epitope
10 (desirably with a high degree of specificity). The fusion protein can usually be eluted from the column by addition of an appropriate buffer.

Additional N-terminal or C-terminal sequences may, however, be present simply as a result of a particular technique used to obtain a polypeptide and need not provide any particular
15 advantageous characteristic to the polypeptide. Such polypeptides are within the scope of the present invention.

Whatever additional N-terminal or C-terminal sequence is present, it is preferred that the resultant polypeptide should exhibit the immunological or biological activity of the
20 polypeptide having the amino acid sequence shown in Figure 1.

Polypeptides within the scope of b)

Turning now to the polypeptides defined in b) above, it will be appreciated by the person
25 skilled in the art that these polypeptides are variants of the polypeptide given in a) above. Such variants preferably exhibit the immunological or biological activity of the polypeptide having the amino acid sequence shown in Figure 1.

Alterations in the amino acid sequence of a protein can occur which do not affect the function of a protein. These include amino acid deletions, insertions and substitutions and can result from alternative splicing and/or the presence of multiple translation start sites and stop sites. Polymorphisms may arise as a result of the infidelity of the translation process. Thus changes in amino acid sequence may be tolerated which do not affect the protein's biological or immunological function.

The skilled person will appreciate that various changes can often be made to the amino acid sequence of a polypeptide which has a particular activity to produce variants (sometimes known as "muteins") having at least a proportion of said activity, and preferably having a substantial proportion of said activity. Such variants of the polypeptides described in a) above are within the scope of the present invention and are discussed in greater detail below. They include allelic and non-allelic variants.

An example of a variant of the present invention is a polypeptide as defined in a) above, apart from the substitution of one or more amino acids with one or more other amino acids. The skilled person is aware that various amino acids have similar properties. One or more such amino acids of a polypeptide can often be substituted by one or more other such amino acids without eliminating a desired activity of that polypeptide.

Thus, the amino acids glycine, alanine, valine, leucine and isoleucine can often be substituted for one another (amino acids having aliphatic side chains). Of these possible substitutions, it is preferred that glycine and alanine are used to substitute for one another (since they have relatively short side chains) and that valine, leucine and isoleucine are used to substitute for one another (since they have larger aliphatic side chains which are hydrophobic).

Other amino acids which can often be substituted for one another include:

- phenylalanine, tyrosine and tryptophan (amino acids having aromatic side chains);

- lysine, arginine and histidine (amino acids having basic side chains);
- aspartate and glutamate (amino acids having acidic side chains);
- asparagine and glutamine (amino acids having amide side chains); and
- cysteine and methionine (amino acids having sulphur-containing side chains).

5

Substitutions of this nature are often referred to as “conservative” or “semi-conservative” amino acid substitutions.

10

Amino acid deletions or insertions may also be made relative to the amino acid sequence given in a) above. Thus, for example, amino acids which do not have a substantial effect on the biological and/or immunological activity of the polypeptide, or at least which do not eliminate such activity, may be deleted. Such deletions can be advantageous since the overall length and the molecular weight of a polypeptide can be reduced whilst still retaining activity. This can enable the amount of polypeptide required for a particular purpose to be reduced – for example, dosage levels can be reduced.

15

Amino acid insertions relative to the sequence given in a) above can also be made. This may be done to alter the properties of a polypeptide used in the present invention (e.g. to assist in identification, purification or expression, as explained above in relation to fusion proteins).

20

Amino acid changes relative to the sequence given in a) above can be made using any suitable technique e.g. by using site-directed mutagenesis (Hutchinson *et al.*, 1978, *J. Biol. Chem.* **253**:6551).

25

It should be appreciated that amino acid substitutions or insertions within the scope of the present invention can be made using naturally occurring or non-naturally occurring amino acids. Whether or not natural or synthetic amino acids are used, it is preferred that only L-amino acids are present.

Whatever amino acid changes are made (whether by means of substitution, insertion or deletion), preferred polypeptides of the present invention have at least 50% sequence identity with a polypeptide as defined in a) above, more preferably the degree of sequence identity is at least 75%. Sequence identities of at least 90% or at least 95% are most preferred.

The percent identity of two amino acid sequences or of two nucleic acid sequences is determined by aligning the sequences for optimal comparison purposes (*e.g.*, gaps can be introduced in the first sequence for best alignment with the sequence) and comparing the amino acid residues or nucleotides at corresponding positions. The “best alignment” is an alignment of two sequences which results in the highest percent identity. The percent identity is determined by the number of identical amino acid residues or nucleotides in the sequences being compared (*i.e.*, % identity = # of identical positions/total # of positions x 100).

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm known to those of skill in the art. An example of a mathematical algorithm for comparing two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. The NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410 have incorporated such an algorithm. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilised as described in Altschul et al. (1997) *Nucleic Acids Res.*

25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molecules (*Id.*). When utilising BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. See <http://www.ncbi.nlm.nih.gov>.

5

Another example of a mathematical algorithm utilised for the comparison of sequences is the algorithm of Myers & Miller, CABIOS (1989). The ALIGN program (version 2.0) which is part of the CGC sequence alignment software package has incorporated such an algorithm. Other algorithms for sequence analysis known in the art include ADVANCE
10 and ADAM as described in Torellis & Robotti (1994) *Comput. Appl. Biosci.*, 10 :3-5; and FASTA described in Pearson & Lipman (1988) *Proc. Natl. Acad. Sci.* 85:2444-8. Within FASTA, ktup is a control option that sets the sensitivity and speed of the search.

Where high degrees of sequence identity are present there will be relatively few differences
15 in amino acid sequence. Thus for example they may be less than 20, less than 10, or even less than 5 differences.

Polypeptides within the scope of c)

20 As discussed *supra*, it is often advantageous to reduce the length of a polypeptide, provided that the resultant reduced length polypeptide still has a desired activity or can give rise to useful antibodies. Feature c) of the present invention therefore covers fragments of polypeptides a) or b) above.

25 The skilled person can determine whether or not a particular fragment has activity using the techniques disclosed above. Preferred fragments are at least 10 amino acids long. They may be at least 20, at least 50 or at least 100 amino acids long.

A polypeptide as defined herein may be useful as antigenic material, and may be used in the production of vaccines for treatment or prophylaxis of cancer. Such material can be "antigenic" and/or "immunogenic". Generally, "antigenic" is taken to mean that the protein is capable of being used to raise antibodies or indeed is capable of inducing an antibody response in a subject. "Immunogenic" is taken to mean that the protein is capable of eliciting a protective immune response in a subject. Thus, in the latter case, the protein may be capable of not only generating an antibody response but, in addition, non-antibody based immune responses.

It is well known that is possible to screen an antigenic protein or polypeptide to identify epitopic regions, i.e. those regions which are responsible for the protein or polypeptide's antigenicity or immunogenicity. Methods well known to the skilled person can be used to test fragments and/or homologues and/or derivatives for antigenicity. Thus, the fragments of the present invention may include one or more such epitopic regions or be sufficiently similar to such regions to retain their antigenic/immunogenic properties. Thus, for fragments according to the present invention the degree of identity is perhaps irrelevant, since they may be 100% identical to a particular part of a protein or polypeptide, homologue or derivative as described herein. The key issue may be that the fragment retains the antigenic/immunogenic properties of the protein from which it is derived.

Homologues, derivatives and fragments may possess at least a degree of the antigenicity/immunogenicity of the protein or polypeptide from which they are derived.

In a further aspect, the present invention provides the use of a polypeptide as defined herein in the production of a medicament for the treatment or prophylaxis of cancer, wherein the medicament is a vaccine. The vaccine optionally comprises one or more suitable adjuvants. Examples of adjuvants well-known in the art include inorganic gels,

such as aluminium hydroxide, and water-in-oil emulsions, such as incomplete Freund's adjuvant. Other useful adjuvants will be well known to the skilled person.

In yet further aspects, the present invention provides:

- 5 (a) the use of a polypeptide as defined herein in the preparation of an immunogenic composition, preferably a vaccine;
- (b) the use of such an immunogenic composition in inducing an immune response in a subject; and
- (c) a method for the treatment or prophylaxis of breast cancer in a subject, or of
10 vaccinating a subject against cancer which comprises the step of administering to the subject an effective amount of a polypeptide as defined herein, preferably as a vaccine.

As will be discussed below, the polypeptides used in the present invention will find use in an immunotherapeutic approach to cancer. The skilled person will appreciate that for the
15 preparation of one or more such polypeptides, the preferred approach will be based on recombinant DNA techniques. In addition, nucleic acid molecules encoding the polypeptides or fragments thereof may be used in their own right. Thus, in a fourth aspect, the invention provides a method of screening for and/or diagnosis of cancer in a subject, which method comprises the step of detecting and/or quantifying the amount of a
20 nucleic acid in a biological sample obtained from said subject, wherein the nucleic acid molecule:

- a) comprises or consists of the DNA sequence shown in Figure 1 or its RNA equivalent;
- b) has a sequence which is complementary to the sequences of a);
- 25 c) has a sequence which codes for the same polypeptide as the sequences of a) or b);
- d) has a sequence which shows substantial identity with any of those of a), b) and c); or

e) a sequence which codes for a derivative or fragment of an amino acid molecule shown in Figure 1.

5 In a fifth aspect, the present invention provides a method for the prophylaxis and/or treatment of cancer in a subject, which comprises administering to said subject a therapeutically effective amount of at least one nucleic acid as defined in the fourth aspect of the invention.

10 In a sixth aspect, the present invention provides the use of at least one nucleic acid as defined in the fourth aspect of the invention in the preparation of a medicament for use in the prophylaxis and/or treatment of cancer.

These nucleic acid molecules are now discussed in greater detail.

15 It is preferred if sequences which show substantial identity with any of those of a), b) and c) have e.g. at least 50%, at least 75% or at least 90% or 95% sequence identity.

20 The polypeptides used in the present invention can be coded for by a large variety of nucleic acid molecules, taking into account the well known degeneracy of the genetic code. All of these molecules are within the scope of the present invention. They can be inserted into vectors and cloned to provide large amounts of DNA or RNA for further study. Suitable vectors may be introduced into host cells to enable the expression of polypeptides used in the present invention using techniques known to the person skilled in the art.

25 The term 'RNA equivalent' when used above indicates that a given RNA molecule has a sequence which is complementary to that of a given DNA molecule, allowing for the fact that in RNA 'U' replaces 'T' in the genetic code. The nucleic acid molecule may be in isolated, recombinant or chemically synthetic form.

Techniques for cloning, expressing and purifying proteins and polypeptides are well known to the skilled person. DNA constructs can readily be generated using methods well known in the art. These techniques are disclosed, for example in J. Sambrook *et al*, *Molecular Cloning 2nd Edition*, Cold Spring Harbour Laboratory Press (1989); in Old & Primrose *Principles of Gene Manipulation* 5th Edition, Blackwell Scientific Publications (1994); and in Stryer, *Biochemistry* 4th Edition, W H Freeman and Company (1995). Modifications of DNA constructs and the proteins expressed such as the addition of promoters, enhancers, signal sequences, leader sequences, translation start and stop signals and DNA stability controlling regions, or the addition of fusion partners may then be facilitated.

Normally the DNA construct will be inserted into a vector, which may be of phage or plasmid origin. Expression of the protein is achieved by the transformation or transfection of the vector into a host cell which may be of eukaryotic or prokaryotic origin. Such vectors and suitable host cells form aspects of the present invention.

The nucleotide sequences of the present invention, including DNA and RNA, and comprising a sequence encoding a polypeptide as defined herein (or a fragment, homologue or analogue thereof), may be synthesised using methods known in the art, such as using conventional chemical approaches or polymerase chain reaction (PCR) amplification. The nucleotide sequences of the present invention also permit the identification and cloning of the gene encoding a polypeptide as defined herein from any species, for instance by screening cDNA libraries, genomic libraries or expression libraries.

Knowledge of the nucleic acid structure can be used to raise antibodies and for gene therapy. Techniques for this are well-known by those skilled in the art, as discussed in more detail herein.

By using appropriate expression systems, polypeptides of the present invention may be expressed in glycosylated or non-glycosylated form. Non-glycosylated forms can be produced by expression in prokaryotic hosts, such as *E. coli*.

5

Polypeptides comprising N-terminal methionine may be produced using certain expression systems, whilst in others the mature polypeptide will lack this residue.

10

Preferred techniques for cloning, expressing and purifying a polypeptide used in the present invention are summarised below:

15

Polypeptides may be prepared natively or under denaturing conditions and then subsequently refolded. Baculoviral expression vectors include secretory plasmids (such as pACGP67 from Pharmingen), which may have an epitope tag sequence cloned in frame (e.g. myc, V5 or His) to aid detection and allow for subsequent purification of the protein. Mammalian expression vectors may include pCDNA3 and pSecTag (both Invitrogen), and pREP9 and pCEP4 (Invitrogen). *E. coli* systems include the pBad series (His tagged - Invitrogen) or pGex series (Pharmacia).

20

In addition to nucleic acid molecules coding for polypeptides used in the present invention, referred to herein as "coding" nucleic acid molecules, the present invention also includes nucleic acid molecules complementary thereto. Thus, for example, both strands of a double stranded nucleic acid molecule are included within the scope of the present invention (whether or not they are associated with one another). Also included are mRNA molecules and complementary DNA Molecules (e.g. cDNA molecules).

25

Nucleic acid molecules which can hybridise to any of the nucleic acid molecules discussed above are also covered by the present invention. Such nucleic acid molecules are referred to

herein as "hybridising" nucleic acid molecules. Hybridising nucleic acid molecules can be useful as probes or primers, for example.

Desirably such hybridising molecules are at least 10 nucleotides in length and preferably are
5 at least 25 or at least 50 nucleotides in length. The hybridising nucleic acid molecules preferably hybridise to nucleic acids within the scope of (a), (b), (c), (d) or (e) above specifically.

Desirably the hybridising molecules will hybridise to such molecules under stringent
10 hybridisation conditions. One example of stringent hybridisation conditions is where attempted hybridisation is carried out at a temperature of from about 35°C to about 65°C using a salt solution which is about 0.9 molar. However, the skilled person will be able to vary such conditions as appropriate in order to take into account variables such as probe length, base composition, type of ions present, etc. For a high degree of selectivity,
15 relatively stringent conditions are used to form the duplexes, such as low salt or high temperature conditions. As used herein, "highly stringent conditions" means hybridisation to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulphate (SDS), 1 mM EDTA at 65 °C, and washing in 0.1xSSC/0.1% SDS at 68 °C (Ausubel F.M. et al., eds., 1989, Current Protocols in Molecular Biology, Vol. I, Green Publishing Associates, Inc., and
20 John Wiley & Sons, Inc., New York, at p. 2.10.3) For some applications, less stringent conditions for duplex formation are required. As used herein "moderately stringent conditions" means washing in 0.2xSSC/0.1% SDS at 42 °C (Ausubel et al., 1989, *supra*). Hybridisation conditions can also be rendered more stringent by the addition of increasing amounts of formamide, to destabilise the hybrid duplex. Thus, particular hybridisation
25 conditions can be readily manipulated, and will generally be chosen depending on the desired results. In general, convenient hybridisation temperatures in the presence of 50% formamide are: 42 °C for a probe which is 95 to 100% identical to the fragment of a gene encoding a polypeptide as defined herein, 37°C for 90 to 95% identity and 32 °C for 70 to 90% identity.

In the preparation of genomic libraries, DNA fragments are generated, some of which will encode parts or the whole of a polypeptide as defined herein. The DNA may be cleaved at specific sites using various restriction enzymes. Alternatively, one may use
5 DNase in the presence of manganese to fragment the DNA, or the DNA can be physically sheared, as for example, by sonication. The DNA fragments can then be separated according to size by standard techniques, including but not limited to agarose and polyacrylamide gel electrophoresis, column chromatography and sucrose gradient centrifugation. The DNA fragments can then be inserted into suitable vectors, including
10 but not limited to plasmids, cosmids, bacteriophages lambda or T₄, and yeast artificial chromosomes (YACs). (See, for example, Sambrook et al., 1989, *Molecular Cloning*, A Laboratory Manual, 1D Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York; Glover, D.M. (ed.), 1985, *DNA Cloning: A Practical Approach*, MRL Press, Ltd., Oxford, U.K. Vol. I, II; Ausubel F.M. et al., eds., 1989, Current Protocols in
15 Molecular Biology, Vol. I, Green Publishing Associates, Inc., and John Wiley & sons, Inc., New York). The genomic library may be screened by nucleic acid hybridisation to labelled probe (Benton & Davis, 1977, *Science* 196:180; Grunstein & Hogness, 1975, *Proc. Natl. Acad. Sci. U.S.A.* 72:3961).

20 Manipulation of the DNA encoding a protein is a particularly powerful technique for both modifying proteins and for generating large quantities of protein for purification purposes. This may involve the use of PCR techniques to amplify a desired nucleic acid sequence. Thus the sequence data provided herein can be used to design primers for use in PCR so that a desired sequence can be targeted and then amplified to a high degree.

25 Typically primers will be at least five nucleotides long and will generally be at least ten nucleotides long (e.g. fifteen to twenty-five nucleotides long). In some cases, primers of at least thirty or at least thirty-five nucleotides in length may be used.

As a further alternative chemical synthesis may be used. This may be automated. Relatively short sequences may be chemically synthesised and ligated together to provide a longer sequence.

5

In addition to being used as primers and/or probes, hybridising nucleic acid molecules of the present invention can be used as anti-sense molecules to alter the expression of substances of the present invention by binding to complementary nucleic acid molecules. This technique can be used in anti-sense therapy.

10

In a specific embodiment, expression of a polypeptide as defined herein is inhibited by use of antisense nucleic acids. The present invention provides the therapeutic or prophylactic use of nucleic acids comprising at least six nucleotides that are antisense to a gene or cDNA encoding a polypeptide as defined herein or a portion thereof. As used herein, an "antisense" nucleic acid refers to a nucleic acid capable of hybridising by virtue of some sequence complementarity to a portion of an RNA (preferably mRNA) encoding a polypeptide as defined herein. The antisense nucleic acid may be complementary to a coding and/or non-coding region of a mRNA encoding such a polypeptide. Such antisense nucleic acids have utility as compounds that inhibit expression, and can be used in the treatment or prevention of cancer.

20

A hybridising nucleic acid molecule of the present invention may have a high degree of sequence identity along its length with a nucleic acid molecule within the scope of (a)-(e) above (e.g. at least 50%, at least 75% or at least 90% or 95% sequence identity). As will be appreciated by the skilled person, the higher the sequence identity a given single stranded nucleic acid molecule has with another nucleic acid molecule, the greater the likelihood that it will hybridise to a nucleic acid molecule which is complementary to that other nucleic acid molecule under appropriate conditions.

25

In view of the foregoing description the skilled person will appreciate that a large number of nucleic acids are within the scope of the present invention. Unless the context indicates otherwise, nucleic acid molecules of the present invention may have one or more of the following characteristics:

- 1) they may be DNA or RNA;
- 2) they may be single or double stranded;
- 3) they may be provided in recombinant form, e.g. covalently linked to a 5' and/or a 3' flanking sequence to provide a molecule which does not occur in nature;
- 4) they may be provided without 5' and/or 3' flanking sequences which normally occur in nature;
- 5) they may be provided in substantially pure form. Thus they may be provided in a form which is substantially free from contaminating proteins and/or from other nucleic acids; and
- 6) they may be provided with introns or without introns (e.g. as cDNA).

If desired, a gene encoding a polypeptide as defined herein, a related gene, or related nucleic acid sequences or subsequences, including complementary sequences, can also be used in hybridisation assays. A nucleotide encoding a polypeptide as defined herein, or subsequences thereof comprising at least 8 nucleotides, can be used as a hybridisation probe. Hybridisation assays can be used for detection, prognosis, diagnosis, or monitoring of conditions, disorders, or disease states, associated with aberrant expression of genes encoding a polypeptide as defined herein, or for differential diagnosis of patients with signs or symptoms suggestive of breast cancer. In particular, such a hybridisation assay can be carried out by a method comprising contacting a patient sample containing nucleic acid with a nucleic acid probe capable of hybridising to a DNA or RNA that encodes a polypeptide as defined herein, under conditions such that hybridisation can occur, and

detecting or measuring any resulting hybridisation. Nucleotides can be used for therapy of patients having cancer, as described below.

In another embodiment, a preparation of oligonucleotides comprising 10 or more
5 consecutive nucleotides complementary to a nucleotide sequence encoding a polypeptide as defined herein or fragment thereof for use as vaccines for the treatment of cancer. Such preparations may include adjuvants or other vehicles.

In a specific embodiment, nucleic acids comprising a sequence encoding a polypeptide as
10 defined herein or functional derivative thereof, are administered to promote polypeptide function by way of gene therapy. Gene therapy refers to administration to a subject of an expressed or expressible nucleic acid. In this embodiment, the nucleic acid produces its encoded protein that mediates a therapeutic effect by promoting polypeptide function. Any of the methods for gene therapy available in the art can be used according to the
15 present invention.

In a preferred aspect, the compound comprises a nucleic acid as defined herein, such as a nucleic acid encoding a polypeptide as defined herein or fragment or chimeric protein thereof, said nucleic acid being part of an expression vector that expresses a polypeptide as
20 defined herein or fragment or chimeric protein thereof in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the polypeptide coding region, said promoter being inducible or constitutive (and, optionally, tissue-specific). In another particular embodiment, a nucleic acid molecule is used in which the coding sequences and any other desired sequences are flanked by regions that promote homologous
25 recombination at a desired site in the genome, thus providing for intrachromosomal expression of the nucleic acid (Koller & Smithies, 1989, *Proc. Natl. Acad. Sci. USA* 86:8932-8935; Zijlstra *et al.*, 1989, *Nature* 342:435-438).

Delivery of the nucleic acid into a patient may be direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector; this approach is known as *in vivo* gene therapy. Alternatively, delivery of the nucleic acid into the patient may be indirect, in which case cells are first transformed with the nucleic acid *in vitro* and then transplanted into the patient; this approach is known as *ex vivo* gene therapy.

A convenient means for detecting/quantifying the polypeptides used in the present invention involves the use of antibodies. Thus, the polypeptides used in the invention also find use in raising antibodies.

10

In a seventh aspect, the present invention provides the use of an antibody which binds to at least one polypeptide as defined in the first aspect of the invention for screening for and/or diagnosis of breast cancer in a subject. Preferably, the antibody is used for detecting and/or quantifying the amount of a polypeptide as defined in the first aspect of the invention in a biological sample obtained from said subject.

15

In one embodiment, binding of antibody in tissue sections can be used to detect aberrant polypeptide localisation or an aberrant level of polypeptide. In a specific embodiment, antibody to a polypeptide as defined herein can be used to assay a patient tissue (*e.g.*, a breast biopsy) for the level of the polypeptide where an aberrant level of polypeptide is indicative of cancer. As used herein, an "aberrant level" means a level that is increased or decreased compared with the level in a subject free from cancer or a reference level. If desired, the comparison can be performed with a matched sample from the same subject, taken from a portion of the body not affected by cancer.

20

Suitable immunoassays include, without limitation, competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation

25

assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays and protein A immunoassays.

- 5 In an eighth aspect, the present invention provides a method for the prophylaxis and/or treatment of cancer in a subject, which comprises administering to said subject a therapeutically effective amount of an antibody which binds to at least one polypeptide as defined in the first aspect of the invention.
- 10 In a ninth aspect, the present invention provides the use of an antibody which binds to at least one polypeptide as defined in the first aspect of the invention in the preparation of a medicament for use in the prophylaxis and/or treatment of cancer.

Preferred antibodies bind specifically to polypeptides of the present invention so that they
15 can be used to purify and/or inhibit the activity of such polypeptides. The antibodies may be monoclonal or polyclonal.

Thus, the polypeptide used in the invention, its fragments or other derivatives, or analogues thereof, may be used as an immunogen to generate antibodies which
20 immunospecifically bind such an immunogen. Antibodies of the invention include, but are not limited to polyclonal, monoclonal, bispecific, humanised or chimeric antibodies, single chain antibodies, Fab fragments and F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. The term "antibody" as used herein refers to immunoglobulin molecules
25 and immunologically-active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that specifically binds an antigen. The immunoglobulin molecules of the invention can be of any class (e.g., IgG, IgE, IgM, IgD and IgA) or subclass of immunoglobulin molecule.

In the production of antibodies, screening for the desired antibody can be accomplished by techniques known in the art, e.g. ELISA (enzyme-linked immunosorbent assay). For example, to select antibodies which recognise a specific domain of a polypeptide used in the invention, one may assay generated hybridomas for a product which binds to a polypeptide fragment containing such domain. For selection of an antibody that specifically binds a first polypeptide homologue but which does not specifically bind to (or binds less avidly to) a second polypeptide homologue, one can select on the basis of positive binding to the first polypeptide homologue and a lack of binding to (or reduced binding to) the second polypeptide homologue.

For preparation of monoclonal antibodies (mAbs) directed toward a polypeptide used in the invention, any technique which provides for the production of antibody molecules by continuous cell lines in culture may be used. For example, the hybridoma technique originally developed by Kohler and Milstein (1975, *Nature* 256:495-497), as well as the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, 1983, *Immunology Today* 4:72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole *et al.*, 1985, in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAbs used in the invention may be cultivated *in vitro* or *in vivo*. In an additional embodiment of the invention, monoclonal antibodies can be produced in germ-free animals utilising known technology (PCT/US90/02545).

The monoclonal antibodies include but are not limited to human monoclonal antibodies and chimeric monoclonal antibodies (e.g., human-mouse chimeras). A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a human immunoglobulin constant region and a variable region derived from

a murine mAb. (See, e.g., U.S. Patent No. 4,816,567; and U.S. Patent No. 4,816,397)

Humanised antibodies are antibody molecules from non-human species having one or more complementarity determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, e.g., U.S. Patent No. 5,585,089).

Chimeric and humanised monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in WO 87/02671; EP-A-184,187; EP-A-171,496; EP-A-173,494; WO 86/01533; U.S. Patent No. 4,816,567; EP-A-125,023; Better *et al.*, 1988, *Science* **240**:1041-1043; Liu *et al.*, 1987, *Proc. Natl. Acad. Sci. USA* **84**:3439-3443; Liu *et al.*, 1987, *J. Immunol.* **139**:3521-3526; Sun *et al.*, 1987, *Proc. Natl. Acad. Sci. USA* **84**:214-218; Nishimura *et al.*, 1987, *Canc. Res.* **47**:999-1005; Wood *et al.*, 1985, *Nature* **314**:446-449; Shaw *et al.*, 1988, *J. Natl. Cancer Inst.* **80**:1553-1559; Morrison, 1985, *Science* **229**:1202-1207; Oi *et al.*, 1986, *Bio/Techniques* **4**:214; U.S. Patent 5,225,539; Jones *et al.*, 1986, *Nature* **321**:552-525; Verhoeyan *et al.* (1988) *Science* **239**:1534; and Beidler *et al.*, 1988, *J. Immunol.* **141**:4053-4060.

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chain genes, but which can express human heavy and light chain genes. The transgenic mice are immunised in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide used in the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harboured by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, see

Lonberg & Huszar (1995), *Int. Rev. Immunol.* **13**:65-93. For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA) and Genpharm (San Jose, CA) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognise a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognising the same epitope. (Jespers *et al.* (1994) *Bio/technology* **12**:899-903).

The antibodies used in the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular, such phage can be utilised to display antigen binding domains expressed from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, e.g., using labelled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulphide stabilised Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Phage display methods that can be used to make the antibodies used in the present invention include those disclosed in Brinkman *et al.*, *J. Immunol. Methods* **182**: 41-50 (1995); Ames *et al.*, *J. Immunol. Methods* **184**:177-186 (1995); Kettleborough *et al.*, *Eur. J. Immunol.* **24**:952-958 (1994); Persic *et al.*, *Gene* **187** 9-18

(1997); Burton *et al.*, *Advances in Immunology* **57**:191-280 (1994); PCT/GB91/01134; WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Patent Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108.

As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')₂ fragments can also be employed using methods known in the art such as those disclosed in WO 92/22324; Mullinax *et al.*, *BioTechniques* **12**(6):864-869 (1992); and Sawai *et al.*, *AJRI* **34**:26-34 (1995); and Better *et al.*, *Science* **240**:1041-1043 (1988).

Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Patents 4,946,778 and 5,258,498; Huston *et al.*, *Methods in Enzymology* **203**:46-88 (1991); Shu *et al.*, *PNAS* **90**:7995-7999 (1993); and Skerra *et al.*, *Science* **240**:1038-1040 (1988).

The invention further provides for the use of bispecific antibodies, which can be made by methods known in the art. Traditional production of full length bispecific antibodies is based on the coexpression of two immunoglobulin heavy chain-light chain pairs, where the two chains have different specificities (Milstein *et al.*, 1983, *Nature* **305**:537-539). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of 10 different antibody molecules, of which only one has the correct bispecific structure. Purification of the correct molecule, which is usually done by affinity chromatography steps, is rather cumbersome,

and the product yields are low. Similar procedures are disclosed in WO 93/08829, and in Traunecker *et al.*, 1991, *EMBO J.* 10:3655-3659.

According to a different and more preferred approach, antibody variable domains with the
5 desired binding specificities (antibody-antigen combining sites) are fused to
immunoglobulin constant domain sequences. The fusion preferably is with an
immunoglobulin heavy chain constant domain, comprising at least part of the hinge, CH2,
and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1)
containing the site necessary for light chain binding, present in at least one of the fusions.
10 DNAs encoding the immunoglobulin heavy chain fusions and, if desired, the
immunoglobulin light chain, are inserted into separate expression vectors, and are
co-transfected into a suitable host organism. This provides for great flexibility in
adjusting the mutual proportions of the three polypeptide fragments in embodiments when
unequal ratios of the three polypeptide chains used in the construction provide the
15 optimum yields. It is, however, possible to insert the coding sequences for two or all three
polypeptide chains in one expression vector when the expression of at least two
polypeptide chains in equal ratios results in high yields or when the ratios are of no
particular significance.

20 In a preferred embodiment of this approach, the bispecific antibodies are composed of a
hybrid immunoglobulin heavy chain with a first binding specificity in one arm, and a
hybrid immunoglobulin heavy chain-light chain pair (providing a second binding
specificity) in the other arm. It was found that this asymmetric structure facilitates the
separation of the desired bispecific compound from unwanted immunoglobulin chain
25 combinations, as the presence of an immunoglobulin light chain in only one half of the
bispecific molecule provides for a facile way of separation. This approach is disclosed in
WO 94/04690. For further details for generating bispecific antibodies see, for example,
Suresh *et al.*, *Methods in Enzymology*, 1986, 121:210.

The invention provides for the use of functionally-active fragments, derivatives or analogues of the anti-polypeptide immunoglobulin molecules. "Functionally-active" means that the fragment, derivative or analogue is able to elicit anti-anti-idiotypic antibodies (i.e., tertiary antibodies) that recognise the same antigen that is recognised by the antibody from which the fragment, derivative or analogue is derived. Specifically, in a preferred embodiment, the antigenicity of the idiotype of the immunoglobulin molecule may be enhanced by deletion of framework and CDR sequences that are C-terminal to the CDR sequence that specifically recognises the antigen. To determine which CDR sequences bind the antigen, synthetic peptides containing the CDR sequences can be used in binding assays with the antigen by any binding assay method known in the art.

The present invention provides antibody fragments such as, but not limited to, F(ab')₂ fragments and Fab fragments. Antibody fragments which recognise specific epitopes may be generated by known techniques. F(ab')₂ fragments consist of the variable region, the light chain constant region and the CH1 domain of the heavy chain and are generated by pepsin digestion of the antibody molecule. Fab fragments are generated by reducing the disulphide bridges of the F(ab')₂ fragments. The invention also provides heavy chain and light chain dimmers of the antibodies of the invention, or any minimal fragment thereof such as Fvs or single chain antibodies (SCAs) (e.g., as described in U.S. Patent 4,946,778; Bird, 1988, *Science* **242**:423-42; Huston *et al.*, 1988, *Proc. Natl. Acad. Sci. USA* **85**:5879-5883; and Ward *et al.*, 1989, *Nature* **334**:544-54), or any other molecule with the same specificity as the antibody of the invention. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide. Techniques for the assembly of functional Fv fragments in *E. coli* may be used (Skerra *et al.*, 1988, *Science* **242**:1038-1041).

In other embodiments, the invention provides fusion proteins of the immunoglobulins of the invention (or functionally active fragments thereof), for example in which the immunoglobulin is fused via a covalent bond (e.g., a peptide bond), at either the N-terminus or the C-terminus to an amino acid sequence of another protein (or portion thereof, preferably at least 10, 20 or 50 amino acid portion of the protein) that is not the immunoglobulin. Preferably the immunoglobulin, or fragment thereof, is covalently linked to the other protein at the N-terminus of the constant domain. As stated above, such fusion proteins may facilitate purification, increase half-life *in vivo*, and enhance the delivery of an antigen across an epithelial barrier to the immune system.

10

The immunoglobulins used in the invention include analogues and derivatives that are either modified, i.e., by the covalent attachment of any type of molecule as long as such covalent attachment that does not impair immunospecific binding. For example, but not by way of limitation, the derivatives and analogues of the immunoglobulins include those that have been further modified, e.g., by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatisation by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, etc. Additionally, the analogue or derivative may contain one or more non-classical amino acids.

15
20

The foregoing antibodies can be used in methods known in the art relating to the localisation and activity of the polypeptides used in the invention, e.g., for imaging or radioimaging these proteins, measuring levels thereof in appropriate physiological samples, in diagnostic methods, etc. and for radiotherapy.

25

The antibodies of the invention can be produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or by recombinant expression, and are preferably produced by recombinant expression technique.

- 5 Recombinant expression of antibodies, or fragments, derivatives or analogues thereof, requires construction of a nucleic acid that encodes the antibody. If the nucleotide sequence of the antibody is known, a nucleic acid encoding the antibody may be assembled from chemically synthesised oligonucleotides (e.g., as described in Kutmeier *et al.*, 1994, *BioTechniques* 17:242), which, briefly, involves the synthesis of overlapping
10 oligonucleotides containing portions of the sequence encoding antibody, annealing and ligation of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

- Alternatively, the nucleic acid encoding the antibody may be obtained by cloning the
15 antibody. If a clone containing the nucleic acid encoding the particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the antibody may be obtained from a suitable source (e.g., an antibody cDNA library, or cDNA library generated from any tissue or cells expressing the antibody) by PCR amplification using synthetic primers hybridisable to the 3' and 5' ends of the sequence or
20 by cloning using an oligonucleotide probe specific for the particular gene sequence.

- If an antibody molecule that specifically recognises a particular antigen is not available (or a source for a cDNA library for cloning a nucleic acid encoding such an antibody), antibodies specific for a particular antigen may be generated by any method known in the
25 art, for example, by immunising an animal, such as a rabbit, to generate polyclonal antibodies or, more preferably, by generating monoclonal antibodies. Alternatively, a clone encoding at least the Fab portion of the antibody may be obtained by screening Fab expression libraries (e.g., as described in Huse *et al.*, 1989, *Science* 246:1275-1281) for

clones of Fab fragments that bind the specific antigen or by screening antibody libraries (See, e.g., Clackson *et al.*, 1991, *Nature* **352**:624; Hane *et al.*, 1997 *Proc. Natl. Acad. Sci. USA* **94**:4937).

5 Once a nucleic acid encoding at least the variable domain of the antibody molecule is obtained, it may be introduced into a vector containing the nucleotide sequence encoding the constant region of the antibody molecule (see, e.g., WO 86/05807; WO 89/01036; and U.S. Patent No. 5,122,464). Vectors containing the complete light or heavy chain for co-expression with the nucleic acid to allow the expression of a complete antibody molecule
10 are also available. Then, the nucleic acid encoding the antibody can be used to introduce the nucleotide substitution(s) or deletion(s) necessary to substitute (or delete) the one or more variable region cysteine residues participating in an intrachain disulphide bond with an amino acid residue that does not contain a sulphhydryl group. Such modifications can be carried out by any method known in the art for the introduction of specific mutations or
15 deletions in a nucleotide sequence, for example, but not limited to, chemical mutagenesis, *in vitro* site directed mutagenesis (Hutchinson *et al.*, 1978, *J. Biol. Chem.* **253**:6551), PCR based methods, etc.

In addition, techniques developed for the production of "chimeric antibodies" (Morrison *et al.*, 1984, *Proc. Natl. Acad. Sci.* **81**:851-855; Neuberger *et al.*, 1984, *Nature* **312**:604-608; Takeda *et al.*, 1985, *Nature* **314**:452-454) by splicing genes from a mouse antibody
20 molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described *supra*, a chimeric antibody is a molecule in which different portions are derived from different animal
25 species, such as those having a variable region derived from a murine mAb and a human antibody constant region, e.g., humanised antibodies.

Once a nucleic acid encoding an antibody molecule has been obtained, the vector for the production of the antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing the polypeptides used in the invention by expressing nucleic acid containing the antibody molecule sequences are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing an antibody molecule coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. See, for example, the techniques described in Sambrook *et al.* (1990, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and Ausubel *et al.* (eds., 1998, Current Protocols in Molecular Biology, John Wiley & Sons, NY).

The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention.

The host cells used to express a recombinant antibody of the invention may be either bacterial cells such as *Escherichia coli*, or, preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule. In particular, mammalian cells such as Chinese hamster ovary cells (CHO), in conjunction with a vector such as the major intermediate early gene promoter element from human cytomegalovirus is an effective expression system for antibodies (Foecking *et al.*, 198, *Gene* 45:101; Cockett *et al.*, 1990, *Bio/Technology* 8:2).

A variety of host-expression vector systems may be utilised to express an antibody molecule of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also

represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express the antibody molecule of the invention *in situ*. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing antibody coding sequences; yeast (e.g., *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing the antibody coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing antibody coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3 cells) harbouring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter).

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions comprising an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther *et al.*, 1983, *EMBO J.* 2:1791), in which the antibody coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, 1985, *Nucleic Acids Res.* 13:3101-3109; Van Heeke & Schuster, 1989, *J. Biol. Chem.* 264:5503-5509); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can

easily be purified from lysed cells by adsorption and binding to a matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

5

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The antibody coding sequence may be cloned individually into non-essential regions (for example, the polyhedrin gene) of the virus and placed under control of an AcNPV
10 promoter (for example, the polyhedrin promoter). In mammalian host cells, a number of viral-based expression systems (e.g., an adenovirus expression system) may be utilised.

As discussed above, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion
15 desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein.

For long-term, high-yield production of recombinant antibodies, stable expression is preferred. For example, cells lines that stably express an antibody of interest can be
20 produced by transfecting the cells with an expression vector comprising the nucleotide sequence of the antibody and the nucleotide sequence of a selectable (e.g., neomycin or hygromycin), and selecting for expression of the selectable marker. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

25

The expression levels of the antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning,

Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase
5 (Crouse *et al.*, 1983, *Mol. Cell. Biol.* **3**:257).

The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers
10 which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, 1986, *Nature* **322**:52; Kohler, 1980, *Proc. Natl. Acad. Sci. USA* **77**:2197). The coding sequences for the heavy and light chains may
15 comprise cDNA or genomic DNA.

Once the antibody molecule used in the invention has been recombinantly expressed, it may be purified by any method known in the art for purification of an antibody molecule, for example, by chromatography (e.g., ion exchange chromatography, affinity
20 chromatography such as with protein A or specific antigen, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins.

Alternatively, any fusion protein may be readily purified by utilising an antibody specific
25 for the fusion protein being expressed. For example, a system described by Janknecht *et al.* allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht *et al.*, 1991, *Proc. Natl. Acad. Sci. USA* **88**:8972-897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that

the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni^{2+} nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively
5 eluted with imidazole-containing buffers.

In a preferred embodiment, antibodies of the invention or fragments thereof are conjugated to a diagnostic or therapeutic moiety. The antibodies can be used for diagnosis or to determine the efficacy of a given treatment regimen. Detection can be facilitated by
10 coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive nuclides, positron emitting metals (for use in positron emission tomography), and nonradioactive paramagnetic metal ions. See generally U.S. Patent No. 4,741,900 for metal ions which can be conjugated to antibodies
15 for use as diagnostics according to the present invention. Suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; suitable prosthetic groups include streptavidin, avidin and biotin; suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride and phycoerythrin; suitable
20 luminescent materials include luminol; suitable bioluminescent materials include luciferase, luciferin, and aequorin; and suitable radioactive nuclides include ^{125}I , ^{131}I , ^{111}In and ^{99}Tc .

Antibodies used in the invention or fragments thereof can be conjugated to a therapeutic
25 agent or drug moiety to modify a given biological response. The therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A,

pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, α -interferon, β -interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, a thrombotic agent or an anti-angiogenic agent, e.g., angiostatin or endostatin; or, a biological response modifier such as a lymphokine, interleukin-1 (IL-1),
 5 interleukin-2 (IL-2), interleukin-6 (IL-6), granulocyte macrophage colony stimulating factor (GM-CSF), granulocyte colony stimulating factor (G-CSF), nerve growth factor (NGF) or other growth factor.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see,
 10 e.g., Arnon *et al.*, "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld *et al.* (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom *et al.*, "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson *et al.* (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A
 15 Review", in Monoclonal Antibodies '84: Biological And Clinical Applications; Pinchera *et al.* (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabelled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin *et al.* (eds.), pp. 303-16 (Academic Press 1985), and Thorpe *et al.*, "The Preparation And Cytotoxic Properties Of
 20 Antibody-Toxin Conjugates", *Immunol. Rev.*, 62:119-58 (1982).

Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described in U.S. Patent No. 4,676,980.

25 An antibody with or without a therapeutic moiety conjugated to it can be used as a therapeutic that is administered alone or in combination with cytotoxic factor(s) and/or cytokine(s).

As discussed herein, certain polypeptides, nucleic acid molecules and antibodies find use in the treatment or prophylaxis of cancer. Thus, in a tenth aspect, the present invention provides a pharmaceutical formulation comprising at least one polypeptide or fragment thereof, nucleic acid molecule or antibody of the invention, optionally together with one or
5 more pharmaceutically acceptable excipients, carriers or diluents. Preferably, the pharmaceutical formulation is for use as a vaccine and so any additional components will be acceptable for vaccine use. In addition, the skilled person will appreciate that one or more suitable adjuvants may be added to such vaccine preparations.

10 The medicament will usually be supplied as part of a sterile, pharmaceutical composition which will normally include a pharmaceutically acceptable carrier. This pharmaceutical composition may be in any suitable form (depending upon the desired method of administering it to a patient).

15 It may be provided in unit dosage form, will generally be provided in a sealed container and may be provided as part of a kit. Such a kit would normally (although not necessarily) include instructions for use. It may include a plurality of said unit dosage forms.

The pharmaceutical composition may be adapted for administration by any appropriate route,
20 for example by the oral (including buccal or sublingual), rectal, nasal, topical (including buccal, sublingual or transdermal), vaginal or parenteral (including subcutaneous, intramuscular, intravenous or intradermal) route. Such compositions may be prepared by any method known in the art of pharmacy, for example by admixing the active ingredient with the carrier(s) or excipient(s) under sterile conditions.

25

Pharmaceutical compositions adapted for oral administration may be presented as discrete units such as capsules or tablets; as powders or granules; as solutions, syrups or suspensions (in aqueous or non-aqueous liquids; or as edible foams or whips; or as emulsions).

Suitable excipients for tablets or hard gelatine capsules include lactose, maize starch or derivatives thereof, stearic acid or salts thereof.

- 5 Suitable excipients for use with soft gelatine capsules include for example vegetable oils, waxes, fats, semi-solid, or liquid polyols etc.

For the preparation of solutions and syrups, excipients which may be used include for example water, polyols and sugars. For the preparation of suspensions, oils (e.g. vegetable
10 oils) may be used to provide oil-in-water or water in oil suspensions.

Pharmaceutical compositions adapted for transdermal administration may be presented as discrete patches intended to remain in intimate contact with the epidermis of the recipient for a prolonged period of time. For example, the active ingredient may be delivered from the
15 patch by iontophoresis as generally described in *Pharmaceutical Research*, 3(6):318 (1986).

Pharmaceutical compositions adapted for topical administration may be formulated as ointments, creams, suspensions, lotions, powders, solutions, pastes, gels, sprays, aerosols or oils. For infections of the eye or other external tissues, for example mouth and skin, the
20 compositions are preferably applied as a topical ointment or cream. When formulated in an ointment, the active ingredient may be employed with either a paraffinic or a water-miscible ointment base. Alternatively, the active ingredient may be formulated in a cream with an oil-in-water cream base or a water-in-oil base. Pharmaceutical compositions adapted for topical administration to the eye include eye drops wherein the active ingredient is dissolved or
25 suspended in a suitable carrier, especially an aqueous solvent. Pharmaceutical compositions adapted for topical administration in the mouth include lozenges, pastilles and mouth washes.

Pharmaceutical compositions adapted for rectal administration may be presented as suppositories or enemas.

5 Pharmaceutical compositions adapted for nasal administration wherein the carrier is a solid include a coarse powder having a particle size for example in the range 20 to 500 microns which is administered in the manner in which snuff is taken, i.e. by rapid inhalation through the nasal passage from a container of the powder held close up to the nose. Suitable compositions wherein the carrier is a liquid, for administration as a nasal spray or as nasal drops, include aqueous or oil solutions of the active ingredient.

10

Pharmaceutical compositions adapted for administration by inhalation include fine particle dusts or mists which may be generated by means of various types of metered dose pressurised aerosols, nebulisers or insufflators.

15 Pharmaceutical compositions adapted for vaginal administration may be presented as pessaries, tampons, creams, gels, pastes, foams or spray formulations.

Pharmaceutical compositions adapted for parenteral administration include aqueous and non-aqueous sterile injection solution which may contain anti-oxidants, buffers, bacteriostats and solutes which render the formulation substantially isotonic with the blood of the intended recipient; and aqueous and non-aqueous sterile suspensions which may include suspending agents and thickening agents. Excipients which may be used for injectable solutions include water, alcohols, polyols, glycerine and vegetable oils, for example. The compositions may be presented in unit-dose or multi-dose containers, for example sealed ampoules and vials, and may be stored in a freeze-dried (lyophilised) condition requiring only the addition of the sterile liquid carrier, for example water for injections, immediately prior to use.

25

Extemporaneous injection solutions and suspensions may be prepared from sterile powders, granules and tablets.

The pharmaceutical compositions may contain preserving agents, solubilising agents, stabilising agents, wetting agents, emulsifiers, sweeteners, colourants, odourants, salts (substances of the present invention may themselves be provided in the form of a
5 pharmaceutically acceptable salt), buffers, coating agents or antioxidants. They may also contain therapeutically active agents in addition to the substance of the present invention.

Dosages of the polypeptide, nucleic acid or antibody used in of the present invention can vary between wide limits, depending upon the disease or disorder to be treated, the age and
10 condition of the individual to be treated, etc. and a physician will ultimately determine appropriate dosages to be used. This dosage may be repeated as often as appropriate. If side effects develop the amount and/or frequency of the dosage can be reduced, in accordance with normal clinical practice.

15 In view of the importance of DTD in cancer, the following form additional aspects of the present invention:

- i) a method of screening for compounds that modulate, i.e. up-regulate or down-regulate, the expression of a polypeptide of the invention, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide or at
20 least one nucleic acid molecule of the invention in a biological sample;
- ii) a method for monitoring/assessing cancer treatment in a patient, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide, nucleic acid molecule or antibody used in the invention in a biological sample obtained from said patient;
- 25 iii) method for the identification of metastatic cancer cells in a biological sample obtained from a subject, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide, nucleic acid or antibody of the invention;

(iv) methods of treating cancer, comprising administering to a patient a therapeutically effective amount of a compound that modulates (e.g., upregulates or downregulates) or complements the expression or the biological activity (or both) of a polypeptide as defined herein in patients having cancer, in order to (a) prevent the onset or
5 development of cancer; (b) prevent the progression of cancer; or (c) ameliorate the symptoms of cancer.

In the context of the present invention, the biological sample can be obtained from any source, such as a serum sample or a tissue sample, e.g. breast tissue. When looking for
10 evidence of metastasis, one would look at major sites of metastasis such as lymph nodes, liver, lung and/or bone.

In one embodiment, breast or colorectal tissue from a subject is analysed for quantitative detection of a polypeptide as defined herein, wherein a change in abundance of the
15 polypeptide in the breast or colorectal tissue from the subject relative to breast or colorectal tissue from a subject or subjects free from breast or colorectal cancer (*e.g.*, a control sample or a previously determined reference range) indicates the presence of breast or colorectal cancer.

20 The invention also provides diagnostic kits, comprising an antibody against a polypeptide as defined herein. In addition, such a kit may optionally comprise one or more of the following: (1) instructions for using the antibody for diagnosis, prognosis, therapeutic monitoring or any combination of these applications; (2) a labelled binding partner to the antibody; (3) a solid phase (such as a reagent strip) upon which the antibody is
25 immobilised; and (4) a label or insert indicating regulatory approval for diagnostic, prognostic or therapeutic use or any combination thereof. If no labelled binding partner to the antibody is provided, the anti-polypeptide antibody itself can be labelled with a

detectable marker, *e.g.*, a chemiluminescent, enzymatic, fluorescent, or radioactive moiety.

5 The invention also provides a kit comprising a nucleic acid probe capable of hybridising to RNA encoding a polypeptide as defined herein. In a specific embodiment, a kit comprises in one or more containers a pair of primers (*e.g.*, each in the size range of 6-30 nucleotides, more preferably 10-30 nucleotides and still more preferably 10-20 nucleotides) that under appropriate reaction conditions can prime amplification of at least a portion of a nucleic acid encoding a polypeptide as defined herein, such as by
10 polymerase chain reaction (see *e.g.*, Innis *et al.*, 1990, PCR Protocols, Academic Press, Inc., San Diego, CA), ligase chain reaction (see EP 320,308) use of Q β replicase, cyclic probe reaction, or other methods known in the art.

15 The invention provides methods for identifying agents, candidate compounds or test compounds that bind to a polypeptide as defined herein or have a stimulatory or inhibitory effect on the expression or activity of a polypeptide as defined herein.

The present invention also provides assays for use in drug discovery in order to identify or verify the efficacy of compounds for treatment or prevention of cancer. Test compounds
20 can be assayed for their ability to modulate levels of a polypeptide as defined herein in a subject having cancer. Compounds able to modulate levels of a polypeptide as defined herein in a subject having cancer towards levels found in subjects free from cancer or to produce similar changes in experimental animal models of cancer can be used as lead compounds for further drug discovery, or used therapeutically. Expression of a
25 polypeptide as defined herein can be assayed by, for example, immunoassays, gel electrophoresis followed by visualisation, detection of activity, or any other method taught herein or known to those skilled in the art. Such assays can be used to screen candidate

drugs, in clinical monitoring or in drug development, where abundance of a polypeptide as defined herein can serve as a surrogate marker for clinical disease.

5 This invention further provides novel agents identified by the above-described screening assays and uses thereof for treatments as described herein. In addition, the invention also provides the use of an agent which interacts with, or modulates the activity of a polypeptide as defined herein in the manufacture of a medicament for the treatment of cancer.

10 The invention provides for treatment or prevention of various diseases and disorders by administration of a therapeutic compound. Such compounds include but are not limited to: a polypeptide as defined herein and analogues and derivatives (including fragments) thereof; antibodies thereto; nucleic acids encoding a polypeptide as defined herein, analogues, or derivatives; antisense nucleic acids to a gene encoding a polypeptide as
15 defined herein, and agonists and antagonists of a gene encoding a polypeptide as defined herein or agonists and antagonists of a polypeptide as defined herein. An important feature of the present invention is the identification of a gene encoding a polypeptide as defined herein involved in cancer. Cancer can be treated or prevented by administration of a therapeutic compound that modulates function or expression of a polypeptide as
20 defined herein in the tissue of cancer patients.

In one embodiment, one or more antibodies each specifically binding to a polypeptide as defined herein are administered alone or in combination with one or more additional therapeutic compounds or treatments. Examples of such treatments include, but are not
25 limited to, taxol, cyclophosphamide, tamoxifen, and doxorubicin.

The compounds of the invention include but are not limited to any compound, *e.g.*, a small organic molecule, protein, peptide, antibody, nucleic acid, etc. that restores the profile

towards normal with the proviso that such compounds or treatments include, but are not limited to, taxol, cyclophosphamide, tamoxifen, and doxorubicin.

5 In another embodiment, symptoms of cancer may be ameliorated by decreasing the level or activity of a polypeptide as defined herein by using gene sequences encoding a polypeptide as defined herein in conjunction with well-known gene "knock-out," ribozyme or triple helix methods to decrease gene expression of the polypeptide. In this approach, ribozyme or triple helix molecules are used to modulate the activity, expression or synthesis of the gene, and thus to ameliorate the symptoms of cancer. Such molecules
10 may be designed to reduce or inhibit expression of a mutant or non-mutant target gene. Techniques for the production and use of such molecules are well known to those of skill in the art.

15 Endogenous polypeptide expression can also be reduced by inactivating or "knocking out" the gene encoding the polypeptide, or the promoter of such a gene, using targetted homologous recombination (*e.g.*, see Smithies, *et al.*, 1985, *Nature* 317:230-234; Thomas & Capecchi, 1987, *Cell* 51:503-512; Thompson *et al.*, 1989, *Cell* 5:313-321; and Zijlstra *et al.*, 1989, *Nature* 342:435-438). For example, a mutant gene encoding a non-functional polypeptide (or a completely unrelated DNA sequence) flanked by DNA homologous to
20 the endogenous gene (either the coding regions or regulatory regions of the gene encoding the polypeptide) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express the target gene *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the target gene. Such approaches are particularly suited in the agricultural field where
25 modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive target gene (*e.g.*, see Thomas & Capecchi, 1987 and Thompson, 1989, *supra*). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors.

Preferred features of each aspect of the invention are as for each of the other aspects *mutatis mutandis*. The prior art documents mentioned herein are incorporated to the fullest extent permitted by law.

5

Examples

The invention will now be described with reference to the following examples, which should not in any way be construed as limiting the scope of the present invention. The
10 examples refer to the figures in which:

Figure 1: shows the nucleotide and amino acid sequences of DTD.;

Figure 2: shows the alignment of the nucleotide sequence from Figure 1 and of
15 part of the genomic clone with accession AC011406. This illustrates the presence of 2 exons, with a boundary between them localised around bp 724-729 in the sequence from Figure 1. On the left-hand graphs, the diagonal white line represents the alignment between the 2 sequences (the DNA from Figure 1 being shown on the vertical scale and the genomic DNA on the horizontal scale). The
20 cursor shows the position at the beginning of the intron (upper panel) and at the end of the intron (lower panel), and the alignment window at the bottom of the images shows the match between the two sequences before the intron (upper panel) and after the intron (lower panel);

Figure 3: shows tissue distribution of DTD mRNA. Levels of mRNA in normal
25 tissues were quantified by real time RT-PCR. mRNA levels are expressed as the number of copies ng⁻¹ cDNA;

Figure 4: shows the expression of DTD in normal and tumour breast tissues. Levels of DTD mRNA in matched normal and tumour tissues from seven breast cancer patients were measured by real time RT-PCR. mRNA levels are expressed as the number of copies ng^{-1} cDNA; and

5

Figure 5: shows the expression of DTD in normal and tumour breast tissues. Samples 1-25 are tumour samples not involving metastasis to the lymph nodes. Samples 26-50 are tumour samples involving metastases to variable no.s of lymph nodes. The final 3 samples are from normal breast tissue (reduction mammoplasties). mRNA levels are expressed as the number of copies ng^{-1} cDNA. There is a statistically significant difference between all tumour samples and normal samples (T-test, $p < 0.05$).

10

Example 1: Cloning of DTD

15 The gene of DTD has been localised to chromosome 5q31-q34 (Hastbacka J, *et al. Genomics* 1991 Dec;11(4):968-73). A Blast search (<http://www.ncbi.nlm.nih.gov/BLAST/>) with the DTD gene against htgs (High-Throughput Genome Sequences), returns the following GenBank clone: AC011406. The DTD sequence was amplified from colon cDNAs using:

20

Sense 5' aggaagctgaaccatctatctc 3'

Antisense 5' actgggaaatgttgacacttg 3'.

25

It was found that the gene consists of 2 exons, with the boundary between them being around bp724-bp729 in the sequence from Figure 1. Figure 2 illustrates the exon position by showing the alignment of the DNA sequence from Figure 1 with part of the genomic clone AC011406 (the figure was generated with the program Dotlet, accessible at <http://www.isrec.isb-sib.ch/java/dotlet/Dotlet.html>).

Example 2: Expression of DTD mRNA in human tissues

We used real time quantitative RT-PCR (Heid, C.A., *et al. Genome Res.* 6, 986-994
5 (1996); Morrison, T.B., *et al. Biotechniques* 24, 954-958 (1998)) to analyse the
distribution of DTD mRNA in normal human tissues (Fig 2).

Quantification of DTD mRNA by RT-PCR

Real time RT-PCR was used to quantitatively measure DTD expression in normal human
10 tissue mRNAs (Clontech), breast cancer tissues removed during surgery, and normal
breast tissue removed during breast reduction mammoplasty. Ethical approval for the
normal and tumour breast samples was obtained at surgery (University of Oxford, UK).
The primers used for PCR were as follows:

15 sense, 5' ccagtcattgcttattccctg 3',
antisense, 5' gttctcggtcaactgtctcacc 3'.

Reactions containing 10ng cDNA, prepared as described above, SYBR green sequence
detection reagents (PE Biosystems) and sense and antisense primers were assayed on an
20 ABI7700 sequence detection system (PE Biosystems). The PCR conditions were 1 cycle
at 50°C for 2 min, 1 cycle at 95°C for 10 min, and 40 cycles of 95°C for 15s, 65°C for
1min. The accumulation of PCR product was measured in real time as the increase in
SYBR green fluorescence, and the data were analysed using the Sequence Detector
program v1.6.3 (PE Biosystems). Standard curves relating initial template copy number to
25 fluorescence and amplification cycle were generated using the amplified PCR product as a
template, and were used to calculate DTD copy number in each sample.

The distribution of DTD mRNA was restricted to a few tissues, with the highest levels of expression in testis and colon (1000 and 530 copies ng⁻¹ cDNA, respectively), and only low levels of DTD message detected in other normal tissues, including mammary gland.

- 5 The distribution of DTD mRNA in clinical breast carcinomas was measured in matched normal and tumour tissue samples from seven breast cancer patients (Figure 4). DTD expression was in all the tumour samples, relative to their matched normal tissues, with five of the samples showing a 10-fold or more elevation in expression. Thus, DTD shows a restricted pattern of expression in normal human tissues, and is elevated in some breast
10 tumours, suggesting that this protein has potential as a therapeutic target.

- To further examine the expression of this gene in breast cancer tissues, we extended the quantification of DTD mRNA levels to a further 40 tumour samples (Figure 5). DTD expression level was found here again to be significantly correlated with the presence of
15 breast tumour.

CLAIMS

- 5 1. A method of screening for and/or diagnosis of cancer in a subject, which method comprises the step of detecting and/or quantifying the amount of a polypeptide in a biological sample obtained from said subject, wherein the polypeptide:
- a) comprises or consists of the amino acid sequence shown in figure 1;
 - b) is a derivative having one or more amino acid substitutions, deletions or
10 insertions relative to the amino acid sequence shown in figure 1; or
 - c) is a fragment of a polypeptide as defined in a) or b) above, which is at least ten amino acids long.
- 15 2. A method as claimed in claim 1, wherein the polypeptide is provided as part of a fusion polypeptide.
3. A method as claimed in claim 2, wherein the fusion polypeptide comprises Green Fluorescent Protein or the DsRed Fluorescent Protein.
- 20 4. A method for the prophylaxis and/or treatment of cancer in a subject, which comprises administering to said subject a therapeutically effective amount of at least one polypeptide as defined in any one of claims 1 to 3.
- 25 5. The use of at least one polypeptide as defined in any one of claims 1 to 3 in the preparation of a medicament for use in the prophylaxis and/or treatment of cancer.

6. A method of screening for and/or diagnosis of cancer in a subject, which method comprises the step of detecting and/or quantifying the amount of a nucleic acid in a biological sample obtained from said subject, wherein the nucleic acid molecule:
- a) comprises or consists of the DNA sequence shown in Figure 1 or its RNA equivalent;
 - b) has a sequence which is complementary to the sequences of a);
 - c) has a sequence which codes for the same polypeptide as the sequences of a) or b);
 - d) has a sequence which shows substantial identity with any of those of a), b) and c); or
 - e) a sequence which codes for a derivative or fragment of an amino acid molecule shown in Figure 1.
7. A method for the prophylaxis and/or treatment of cancer in a subject, which comprises administering to said subject a therapeutically effective amount of at least one nucleic acid as defined in claim 6.
8. The use of at least one nucleic acid as defined in claim 6 in the preparation of a medicament for use in the prophylaxis and/or treatment of cancer.
9. The use of an antibody which binds to at least one polypeptide as claimed in any one of claims 1 to 3 for screening for and/or diagnosis of cancer in a subject.
10. A method for the prophylaxis and/or treatment of cancer in a subject, which comprises administering to said subject a therapeutically effective amount of an antibody which binds to at least one polypeptide as defined in any one of claims 1 to 3.

11. A method as claimed in claim 10, wherein the antibody which binds specifically to a polypeptide as defined in any one of claims 1 to 3.
12. A method as claimed in claim 10 or claim 11, wherein the antibody is conjugated to a therapeutic moiety.
13. A method as claimed in claim 12, wherein the therapeutic moiety is selected from a second antibody or a fragment or derivative thereof, a cytotoxic agent or a cytokine.
14. The use of an antibody which binds to at least one polypeptide as defined in any one of claims 1 to 3 in the preparation of a medicament for use in the prophylaxis and/or treatment of cancer.
15. The use as claimed in claim 14 modified by the features of any one of claims 11, 12 and 13.
16. The use as claimed in claim 14 or 15, wherein the medicament is a vaccine.
17. A pharmaceutical formulation comprising at least one polypeptide as defined in claim 1, 2 or 3, at least one nucleic acid molecule as defined in claim 6 or at least one antibody as defined in any one of claims 9 to 13, optionally together with one or more pharmaceutically acceptable excipients, carriers or diluents.
18. A pharmaceutical formulation as claimed in claim 17 which is for use as a vaccine.
19. A pharmaceutical formulation as claimed in claim 17 or claim 18, which comprises one or more suitable adjuvants.

20. A method of screening for compounds that modulate, i.e. up-regulate or down-regulate, the expression of a polypeptide as defined in claim 1, 2 or 3, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide as defined in claim 1, 2 or 3, or at least one nucleic acid molecule as defined in claim 6 in
5 a biological sample.
21. A method for monitoring/assessing cancer treatment in a patient, which comprises the step of determining the presence or absence and/or quantifying at least one polypeptide as defined in claim 1, 2 or 3, at least one nucleic acid molecule as defined in claim 6, or at
10 least one antibody as defined in claim 9, 10 or 11 in a biological sample obtained from said patient.
22. A method for the identification of metastatic cancer cells in a biological sample obtained from a subject, which comprises the step of determining the presence or absence
15 and/or quantifying at least one polypeptide as defined in claim 1, 2 or 3, at least one nucleic acid molecule as defined in claim 6, or at least one antibody as defined in claim 9, 10 or 11.

Abstract**Protein**

- 5 The present invention provides methods for the diagnosis, screening, treatment and prophylaxis of cancer using a protein (DTD), nucleic acid encoding the protein and antibody that binds to the protein.

Figure 1.

AGGAAGCTGAACCATCTATCTCCAGAAATGTCTTCAGAAAGTAAAGAGCAACATAACGTT	60
M S S E S K E Q H N V	11
TCACCCAGAGACTCAGCTGAAGGAAATGACAGTTATCCATCTGGGATCCATCTGGAACCTT	120
S P R D S A E G N D S Y P S G I H L E L	31
CAAAGGGAATCAAGTACTGACTTCAAGCAATTTGAGACCAATGATCAATGCAGACCTTAT	180
Q R E S S T D F K Q F E T N D Q C R P Y	51
CATAGGATCCTTATTGAGCGTCAAGAGAAATCAGATACAAACTTCAAGGAGTTGTTATT	240
H R I L I E R Q E K S D T N F K E F V I	71
AAAAAGCTGCAGAAGAATTGCCAGTGCAGTCCAGCCAAAGCCAAAAATATGATTTTAGGT	300
K K L Q K N C Q C S P A K A K N M I L G	91
TTCTTCTCTGTTTTGCAGTGGCTCCCAAAATACGACCTAAAGAAAAACATTTTAGGGGAT	360
F L P V L Q W L P K Y D L K K N I L G D	111
GTGATGTCAGGCTTGATTGTGGGCATATTATTGGTGCCCCAGTCCATTGCTTATCCCTG	420
V M S G L I V G I L L V P Q S I A Y S L	131
CTGGCTGGCCAAGAACCTGTCTATGGTCTGTACACATCTTTTTTGGCAGCATCATTTAT	480
L A G Q E P V Y G L Y T S F F A S I I Y	151
TTTCTCTTGGGTACCTCCCGTCACATCTCTGTGGGCATTTTTTGGAGTACTGTGCCTTATG	540
F L L G T S R H I S V G I F G V L C L M	171
ATTGGTGAGACAGTTGACCGAGAACTACAGAAAGCTGGCTATGACAATGCCCATAGTGCT	600
I G E T V D R E L Q K A G Y D N A H S A	191
CCTTCTTAGGAATGGTTTCAAATGGGAGCACATTATTAATCATAACATCAGACAGGATA	660
P S L G M V S N G S T L L N H T S D R I	211
TGTGACAAAAGTTGCTATGCAATTATGGTTGGCAGCACTGTAACTTTATAGCTGGAGTT	720
C D K S C Y A I M V G S T V T F I A G V	231
TATCAGGTAGCGATGGGCTTCTTTCAAGTGGGTTTTGTTTCTGTCTACCTCTCAGATGCC	780
Y Q V A M G F F Q V G F V S V Y L S D A	251
TTGCTGAGTGGATTGTCACTGGTGCCTCCTTCACTATTCTTACATCTCAGGCCAAGTAT	840
L L S G F V T G A S F T I L T S Q A K Y	271
CTTCTTGGGCTCAACCTTCCTCGGACTAATGGTGTGGGCTCACTCATCACTACCTGGATA	900
L L G L N L P R T N G V G S L I T T W I	291
CATGTCTTCAGAAACATCCATAAGACCAATCTCTGTGATCTTATCACCAGCCTTTTGTGC	960
H V F R N I H K T N L C D L I T S L L C	311
CTTTTGGTTCTTTTGCCAACCAAGAACTCAATGAACACTTCAAATCCAAGCTTAAGGCA	1020
L L V L L P T K E L N E H F K S K L K A	331
CCGATTCCTATTGAACTTGTGTGTGTGTAGCAGCCACATTAGCCTCTCATTTTGAAAA	1080
P I P I E L V V V V A A T L A S H F G K	351
CTACATGAAAATTATAATTCTAGTATTGCTGGACATATTCCCACTGGGTTTATGCCACCC	1140
L H E N Y N S S I A G H I P T G F M P P	371
AAAGTACCAGAATGGAACCTAATTCCTAGTGTGGCTGTAGATGCAATAGCTATTTCCATC	1200
K V P E W N L I P S V A V D A I A I S I	391
ATTGGTTTTGCTATCACTGTATCACTTTCTGAGATGTTTGCCAAGAAACATGGTTACACA	1260
I G F A I T V S L S E M F A K K H G Y T	411

GTCAAAGCAAACCAGGAAATGTATGCCATTGGCTTTTGTAAATATCATCCCTTCCTTCTTC	1320
V K A N Q E M Y A I G F C N I I P S F F	431
CACTGTTTTACTACTAGTGCAGCTCTTGCAAAGACATTGGTTAAAGAATCAACAGGCTGC	1380
H C F T T S A A L A K T L V K E S T G C	451
CATACTCAGCTTTCTGGTGTGGTAACAGCCCTGGTTCTTTTGTGGTCCTCCTAGTAATA	1440
H T Q L S G V V T A L V L L V L L V I	471
GCTCCTTTGTTCTATTCCCTTCAAAAAAGTGCCTTGGTGTGATCACAATTGTAAATCTA	1500
A P L F Y S L Q K S V L G V I T I V N L	491
CGGGGAGCCCTTCGTAAATTTAGGGATCTTCCCAAATGTGGAGTATTAGTAGAATGGAT	1560
R G A L R K F R D L P K M W S I S R M D	511
ACAGTTATCTGGTTTGTACTATGCTGTCTCTGCACTGCTAAGTACTGAAATAGGCCTA	1620
T V I W F V T M L S S A L L S T E I G L	531
CTTGTTGGGGTTTGTCTTCTATATTTTGTGTCTCCTCCGCACTCAGAAGCCAAAGAGT	1680
L V G V C F S I F C V I L R T Q K P K S	551
TCACTGCTTGGCTTGGTGAAGAGTCTGAGGTCTTTGAATCTGTGTCTGCTTACAAGAAC	1740
S L L G L V E E S E V F E S V S A Y K N	571
CTTCAGACTAAGCCAGGCATCAAGATTTTCCGCTTTGTAGCCCTCTCTACTACATAAAC	1800
L Q T K P G I K I F R F V A P L Y Y I N	591
AAAGAATGCTTTAAATCTGCTTTTATACAAACAACTGTCAACCCAATCTTAATAAAGGTG	1860
K E C F K S A L Y K Q T V N P I L I K V	611
GCTTGAAGAAGGCAGCAAAGAGAAAGATCAAAGAAAAAGTAGTGAAGTCTTGGTGAATC	1920
A W K K A A K R K I K E K V V T L G G I	631
CAGGATGAAATGTCAGTGCAACTTTCCCATGATCCCTTGGAGCTGCATACTATAGTGATT	1980
Q D E M S V Q L S H D P L E L H T I V I	651
GACTGCAGTGCAATTCAATTTTTAAATACAGCAGGGATCCACACTGAAAGAAGTTCGC	2040
D C S A I Q F L N T A G I H T L K E V R	671
AGAGATTATGAAGCCATTGGAATCCAGGTTCTGCTGGCTCAGTGCAATCCCACTGTGAGG	2100
R D Y E A I G I Q V L L A Q C N P T V R	691
GATTCCCTAACCAACGGAGAATATTGCAAAAAGGAAGAAGAAAACCTTCTCTTCTATAGT	2160
D S L T N G E Y C K K E E E N L L F Y S	711
GTGTATGAAGCGATGGCTTTTGCAAGATATCTAAAAATCAGAAAGGAGTATGTGTTCCC	2220
V Y E A M A F A E V S K N Q K G V C V P	731
AATGGTCTGAGTCTTAGTAGTGATTAATTGAGAAGGTAGATAGAAGAATGTCTAGCCAAT	2280
N G L S L S S D *	739
AGGTTAGGATTTCAAGTGTCCAACATTTCCAGT	2314

Figure 3

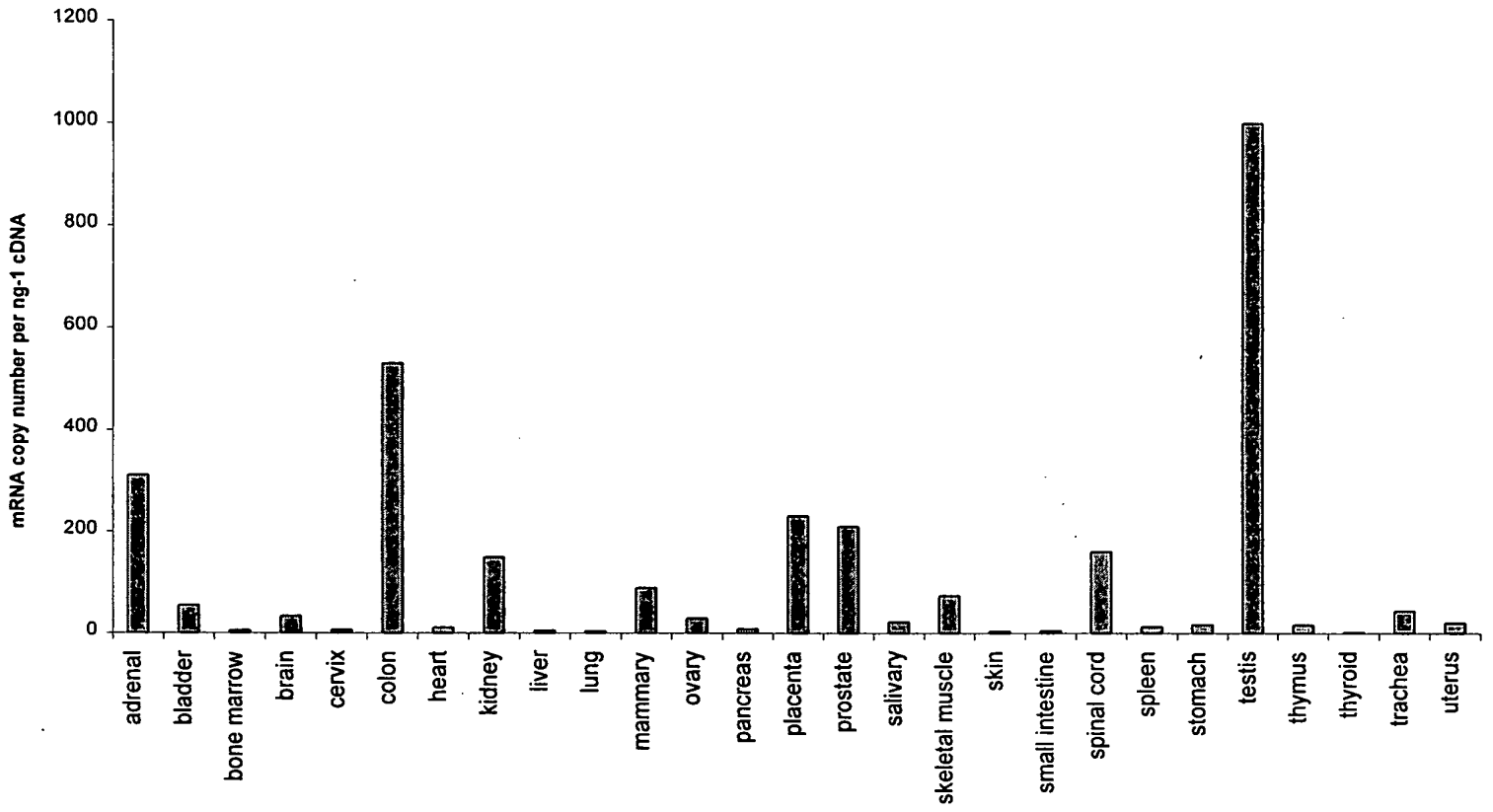


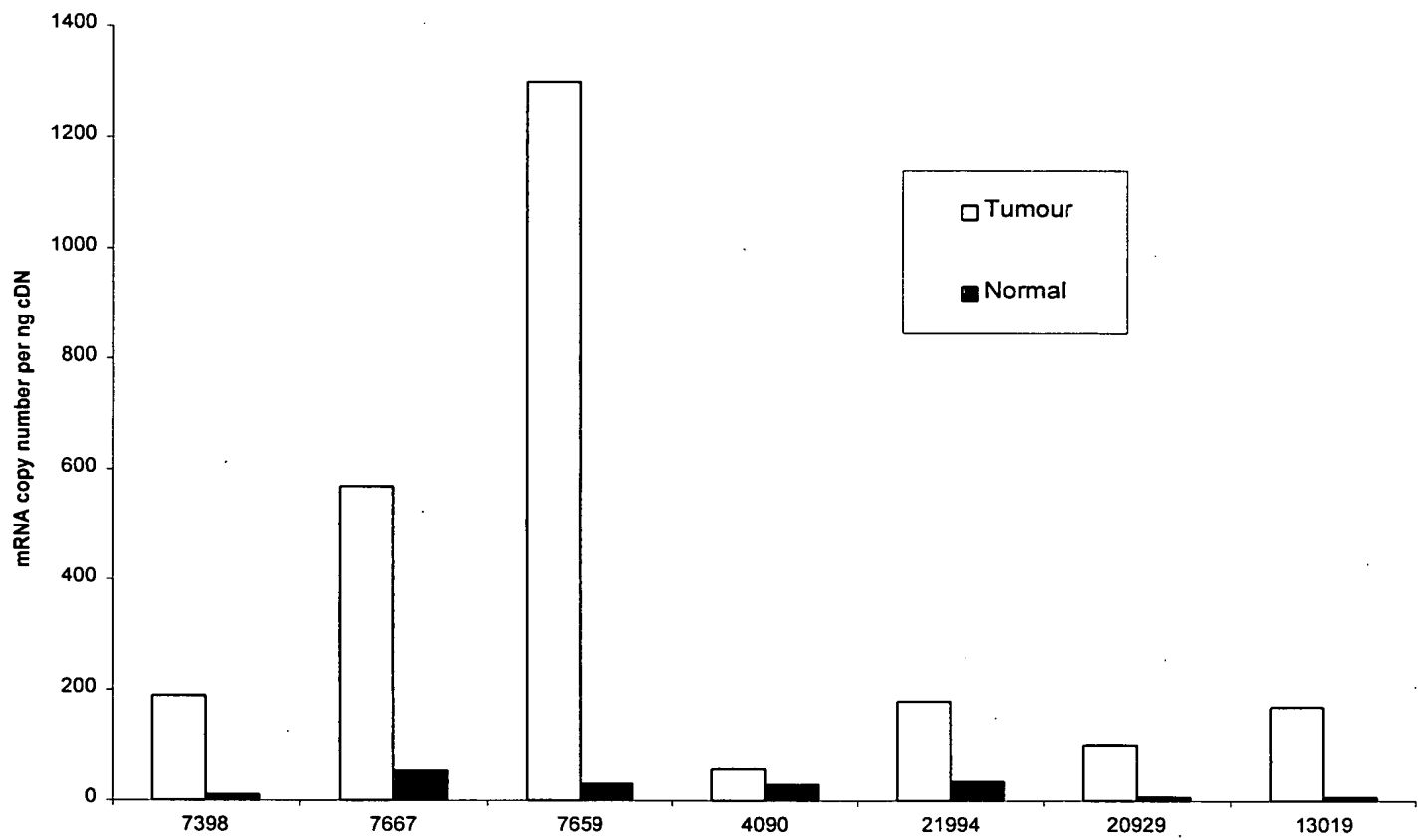
Figure 4

Figure 5

